The Effect of Explanatory Captions on the Reception of Foreign Audiovisual Products: A Study Drawing on Eye-tracking Data and Retrospective Interviews

Binghan ZHENG* and Mingqing XIE

Durham University, UK

*Binghan Zheng (Corresponding author) (ORCID: 0000-0001-5302-4709)
School of Modern Languages & Cultures
Durham University
Elvet Riverside, New Elvet, Durham
DH1 3JT, UK
Email:binghan.zheng@durham.ac.uk

The Effect of Explanatory Captions on the Reception of Foreign Audiovisual Products: A Study Drawing on Eye-tracking Data and Retrospective Interviews

Abstract: The present research triangulates questionnaire, retrospective interview and eye-tracking data, aiming to investigate how Explanatory Captions (ECs) are received by different viewers with varied educational backgrounds, and whether or not the presence of ECs improves their understanding of the AV content. The results show that the provision of ECs, for a subtitled video in a foreign language, greatly increased positive cognitive effects (PCE) on the viewers. Viewers tend to reduce time spent on viewing images, but invest additional processing effort on the ECs, although their allocation of processing effort on subtitles experienced little change. Furthermore, the eye-tracking data suggest that most participants adopted a fixed reading pattern on subtitles and ECs when they appeared simultaneously, which could balance some of the negative impact of ECs on their viewing experience. The findings gained through this experimental research will provide some guidance and suggestions for subtitlers when preparing subtitles and ECs.

Keywords: Explanatory Caption; Reception; Relevance Theory; Eye-tracking; Retrospection

1. Introduction

The past decade has witnessed the rapid growth of fan-sub groups in China. The members are mostly amateur translators and their translations are normally voluntary. With no agreed standard practice for subtitling imposed on their practice, various subtitle formats for foreign TV series and films have emerged, among which the Explanatory Caption[1] (EC) is probably the most significant. EC in this paper is defined as a caption at the top of the screen aimed at improving the viewers' comprehension of the video, especially on imported AV products which can pose many linguistic and cultural barriers (humour, puns, idioms, allusions, etc.) for foreign language audiences. There are mainly two forms of ECs: one is to translate the text appearing in the video that conveys critical information without which the audience

[1] Explanatory Caption as a terminology has been previously coined in the area of information graphics. In this research, however, it has been redefined as a caption aiming at improving the viewers' comprehension of an AV product.

might be confused about the plot; the other is to explain the 'untranslatable' items, either linguistic or non-linguistic, in the video (see Figure 1).

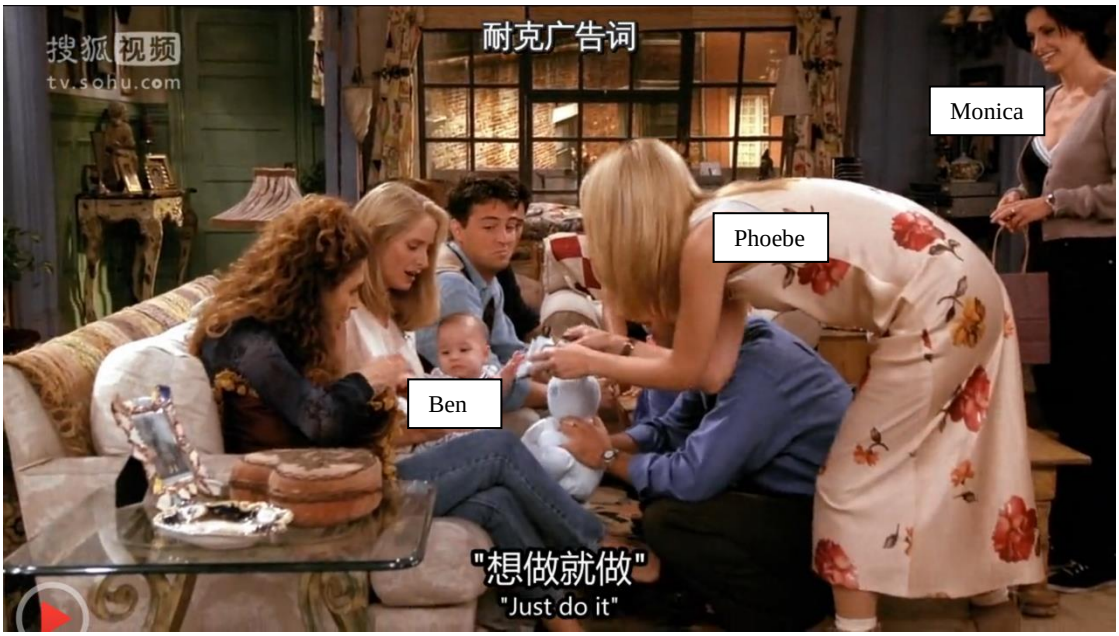

Figure 1. Scene from *Friends* (American television sitcom), Season 2 Episode 02

In the depicted scene Monica bought a pair of Nike shoes for Ben as a gift. When Phoebe gave them to the baby, she said '*Just do it*', which is the Nike slogan, and laughter can be heard in the background. An EC (*耐克广告词*, gloss: "a Nike slogan") is then inserted to explain this, so as to avoid any possible confusion among Chinese audiences who might not be familiar with the slogan.

ECs have become widely accepted among both Chinese subtitlers and online AV users. A typical example is the popularity of a Chinese subtitler, *Gudabaihua* (*谷大白话*), who gained nearly nine million followers on his *Weibo* (a popular social network in China) by clarifying English slang and culturally-specific concepts in American talk-shows and comedy videos using ECs (http://weibo.com/ichthy?profile_ftype=1&is_all=1#_0). The fact that some famous AV websites have started adopting ECs possibly suggests that, ECs will become a common subtitling practice in the near future, not only for online videos, but also for official TV programmes and films.

Despite the growing popularity of ECs for AV products in China, however, there is hardly any research examining their actual effects on improving viewers' reception and comprehension of AV content. With a view to addressing this void, our research aims to investigate how ECs are perceived by foreign language viewers and whether the presence of

ECs does improve their comprehension of AV content. Sperber & Wilson's (1995) Relevance Theory is applied to provide criteria for evaluating the effects of ECs; and eye-tracking, questionnaires and retrospective interviews are used as methods for collecting triangulated data.

2. Theoretical Framework

2.1. Relevance Theory

One of the results of long-term natural selection has been the evolution of a minimisation strategy on the part of the human cognition system: it aims to invest less cognitive effort yet gain greater cognitive effect which, according to the definition of Sperber & Wilson (1995, 265), is "a contextual effect occurring in a cognitive system (e.g., an individual)". In relevance-theoretic terms this is 'maximisation of relevance' (Cognitive Principle of Relevance) (Sperber & Wilson 1995). Based on such a cognitive mechanism, Relevance Theory (RT) argues, in an 'ostensive-inferential communication', a communicator creates an ostensive stimulus (a sound, an utterance, an action, etc.) that is assumed to be of *optimal relevance* to the audience (Communicative Principle of Relevance), to attract an audience's attention. In other words, the audience is led to believe that in order to gain some positive cognitive effects, such a stimulus is relevant enough to be worth processing. Then the audience, constantly evaluating relevance and retrieving data first from *explicatures* and then from *implicatures* contained in the stimulus, infer an interpretation that is sufficiently relevant (Sperber & Wilson 1995; Wilson & Sperber 2002; Braun 2016). It is worth noting that the constant effort-efficient relevance-seeking mechanism does not exclude inputs or interpretations that demand more processing effort, but rather that greater processing effort normally yields greater cognitive effect (Yus 2008; Braun 2016).

Subtitling is a complex type of communication, involving two levels of sub-communication. Firstly, video makers, as the first level communicators, integrate several semiotic modes (e.g., dialogues, images and soundtrack) and create a multimodal text (i.e., a video) as a stimulus that is assumed to be optimally relevant to the first level audiences, who understand the original culture and the language spoken in the video. These include the subtitlers. As the second level communicators, the subtitlers create subtitles as a stimulus, integrated with the other video components, to maximise the relevance of the video to the second level audiences (i.e., foreign language audiences) who are less proficient in the original language and culture. Subtitles, which offer written translations of the dialogues in

the video, aim to compensate for the insufficient relevance of the video to the second level audiences caused by their lack of linguistic and cultural context. As Braun (2016) argues therefore, subtitling poses a challenge to subtitlers in assessing not only the contribution of all the video components, but also the knowledge of the target audiences.

ECs are also a particular type of stimulus created by subtitlers. Differing slightly from subtitles, which are more explicit and direct, ECs aim at explaining items contained in the verbal or non-verbal expressions that are implicit and indirect, especially for a second level audience, such as the wordplay in Figure 1. These implicit and indirect items are deliberately created by the video makers to invite viewers to engage in the meaning-making procedure and to make the programme more interesting and appealing (Tannen 1989; Thomas 1995; Desilla 2012). In other words, without these items, the intended meaning might not be fully expressed and the force of the programme might be weakened. Again, these items require linguistic and cultural background knowledge that foreign language audiences usually lack, thus reducing the relevance of the AV programme. The EC is a technical and practical approach to providing the essential information necessary to compensate and bridge the comprehension gap. Whether ECs are truly of sufficient relevance to audiences should be evaluated by measuring viewers' Positive Cognitive Effect (PCE) and processing efforts (Wilson & Sperber 2002). That is, if the processing efforts remain almost the same with the provision of an EC, but the PCE experienced by viewers is obviously enhanced, it means that the EC is sufficiently relevant to the audience, and thus has improved their comprehension of the AV content.

2.2. Positive cognitive effect (PCEs) and processing effort

Sperber and Wilson (1995, 265) define PCE as "a cognitive effect that contributes positively to the fulfilment of cognitive functions or goals", such as an increase in knowledge. In this research, PCE refers to an improvement in a viewer's understanding of the AV content. Specifically, if the viewers have accurately accessed the information conveyed by the AV programme, it means they have gained some PCEs; if they gain information which deviates from the intended meaning of the AV programme, it means they have gained some cognitive effects, but not positive ones. The following example gives a clear explanation of the difference between PCEs and other cognitive effects.

In the scene depicted in Figure 2, the characters Ross, Joey, and Chandler are making fun of Monica, because she has dated a high school boy. Ross crosses his hands as shown in the picture and shouts: "*It's Morphin time!*". Then, Joey makes the same gesture and shouts:

"*Tigerzord!*", followed by Chandler shouting: "*Tyrannosaurus!*". These lines and gestures are from the then-popular American children's TV series, *Mighty Morphin Power Rangers*. By doing this, Ross, Joey and Chandler are teasing Monica over dating a 'young child'. This is what makes the scene entertaining, and thus loud laughter can be heard from the live audience. Viewers might have several different reactions to the scene, for example:

Figure 2. Scene from *Friends* (American television sitcom), Season 1 Episode 22

(1) some might feel confused about what makes the audiences laugh so loudly;
(2) some might laugh because they find the actions of the three men hilarious;
(3) some might laugh because they know that these three men are mocking Monica again by imitating characters in the *Mighty Morphin Power Rangers*.

Audiences with the first reaction have gained barely any cognitive effect from this scene, because almost no change has occurred in their representation of the world. In the case of those viewers who are amused just by the funny gestures and facial expressions of the actors, even though such laughter is usually considered as positive, the cognitive effects these viewers have gained are hardly positive. Their understanding of the scene has deviated from the intended meaning of the video maker. By contrast, those with the third reaction have successfully grasped the implied deeper meaning of the scene, and gained more fun from this TV series, an outcome which can be regarded as a worthwhile addition to their representation of the world. Only in this case, can the viewers be considered to have gained some PCEs.

Processing effort as defined by Wilson & Sperber (2002, 252) is "the effort of perception, memory and inference" invested in processing information. As a hypothetical conception, processing effort can only be observed through its effects (Caffery 2013, 230). Among all the effects, eye movement has been proven to have a close link with information processing (Rayner 1998; 2009). As a result of development in eye-tracking technology in the past few decades, researchers have been able to investigate the relationship between eye movement and the allocation of attention.

3. Eye-tracking research on subtitle reception

In the past two decades, eye-tracking technology has been widely applied to the investigation of the reception of subtitling, thus aiming to understand how subtitled AV products are received and perceived by viewers (cf. Perego 2012). D'Ydewalle, Van Rensbergen and Pollet (1987) demonstrated that people who understand the spoken language(s) in the video still tend to read subtitles, as subtitles are read automatically and effortlessly (d'Ydewalle & Gielen 1992; Bisson et al. 2014) and it is more efficient to gather information from subtitles than the soundtrack (Moran 2012, 186). Furthermore, in a multimodal text, the processing of text, including subtitles, is often prioritised over that of other semiotic modes (Hegarty 1992). Kruger et al. (2013), by comparing the cognitive loads of viewing educational videos with or without subtitles, also found cognitive loads when watching videos with subtitles were not higher but lower. Eye movements between several signal inputs, such as images and texts, i.e., *deflections* (de Linde & Kay 1999), will possibly reduce the semantic impact of subtitles (Lautenbacher 2013, 148).

Reading subtitles can be influenced by many factors, such as the layout of the subtitle and the number of subtitle lines (cf. d'Ydewalle & Gielen 1992; d'Ydewalle & De Bruycker 2007). The languages of the soundtrack also influence the subtitle reading. Bisson et al. (2014) illustrate a more regular reading of subtitles when the soundtrack is in a foreign language that is unknown to the viewer, compared with when the soundtrack is in the viewer's native language.

Subtitle translation strategies also influence viewers' eye movements. They tend to exhibit more deflections when watching a film with non-literally translated subtitles, than that with literally translated ones (Ghia 2012). Although condensing is widely adopted by subtitlers due to the limited time frame and screen width, Moran (2012) argues that the emphasis on conciseness in subtitling is of debatable benefit. His experimental results show that the usage frequency of the words in subtitles has an inversely proportional effect on the

fixation duration and total gaze time, and that the same is true for the degree of the subtitle cohesion, indicating that it is easier to process more cohesive and longer subtitles.

Being a similar presentation style and purpose to subtitles, ECs have rarely been investigated by researchers of AV translation and subtitling. Dwyer (2015) investigated the perception of on-screen texts, which convey the thought procedure of the character of Sherlock, in the popular BBC TV series *Sherlock* (2010-), and depict the mobile phone messaging. This technique is also used to some extent to draw the audience's attention away from other parts of the image, nevertheless, such distraction and the increase of processing loads are acceptable to the audience. There are two major differences however, between the on-screen text in *Sherlock* and the ECs in the current research: first, while ECs are normally placed statically on the top on the screen, the on-screen text could appear in any part of the screen; secondly, the on-screen text is normally in the language spoken in the video and does not involve any translation, while the ECs in this study are in the translated language and aim at reinterpreting critical on-screen texts or explaining the 'untranslatable' items. More in line with the current research is Caffrey's (2012) research on the "culturally marked visual nonverbal cues (CVNCs)". The function of pop-up glosses relating to CVNCs and ECs are mostly the same: to provide background knowledge of some culture-specific items and thus enhance the target viewer's comprehension. While the position of the EC is relatively fixed, again, pop-up glosses could appear anywhere on the screen. According to Caffrey's research results, pop-up glosses improve viewers' understanding of CVNCs but also increase the processing efforts, and that the presence of pop-up glosses increases the number of subtitles skipped by the viewers (2012, 253-256).

Building on the previous eye-tracking research on subtitle reception, we aim to investigate the following research questions through an empirical-experimental study drawing on eye-tracking data and retrospective interviews: 1) How are the ECs perceived by foreign language viewers? i.e., how do the viewers allocate their attention among subtitles, ECs and images? Will the subtitle or the EC be read in the first instance by viewers? 2) Does the presence of ECs improve the comprehension of the AV content? 3) Is the presence of ECs acceptable for foreign language viewers? i.e., Do ECs increase the cognitive efforts of the viewers? If so, is the amount of increased efforts acceptable?

4. Method

This research adopted a between-subject design with the inclusion of ECs in the stimulus video as the only independent variable. The video without ECs was used for the control group while the one with 19 ECs was shown to the test group. Both groups were asked to complete three formal experimental components: viewing a video presented on the screen of an eye-tracker, answering a questionnaire with questions about themselves and about the video, and taking part in a brief retrospective interview related to their questionnaire answers.

4.1. Participants

A group of 38 volunteers took part in the experiment and each received £5 for their participation. They were all native Chinese speakers, with English as their second language. They all had normal or corrected-to-normal vision. With the exception of the 2 participants undertaking the pilot tests, and 2 participants data being excluded from further analysis due to relatively low quality eye-tracking data, the remaining 34 participants were categorised into three levels based on their educational background (see Table 1). Although there is no IELTS score record for the UL group, they are all Chinese university lecturers teaching English-Chinese translation, and thus assumed to be fully proficient in using English as a foreign language. The categorisation into different levels based on different educational backgrounds reflected an attempt to test a heterogeneous sample of target viewers, but in this study no comparative research was conducted among the three groups. An equal number of participants from each level were randomly assigned to either the control or test group for the experiment.

Table 1. Participants by age and education background

	Pre-sessional Students for MA courses (PS)	**Translation Students for MA courses (TS)**	**University Lecturers in Translation Studies (UL)**
Number	16	12	6
IELTS Score (Mean/SD)	6.41 /0.375	7.38 /0.376	/
Age (Mean/SD)	24.06 /4.82	23.83 /1.27	41/2.58

4.2. Stimuli

The material used in the experiment is a 3:46-minute video clip from the *2016 White House Correspondence Dinner (2016 WHCD)*, in which the then U.S. President Obama envisages his life after retirement. The video was chosen for its relatively low exposure to Chinese

audiences and the numerous phrases and episodes included in the video that allude to political events in the USA. The low exposure reduces the impact on the experimental results caused by participants' previous viewing, while the abundant culture-specific terms and allusions set sufficient challenges for non-English native participants in our experiment.

Two versions of the video were employed: one with both Mandarin Chinese subtitles and ECs, the other with the same subtitles but no ECs. Both the subtitles and ECs were supplied by a professional subtitler who had over 5 years working experience on subtitling. The 'foreignisation' strategy was applied to reduce deflections (Ghia 2012) and to simulate common practice in China online subtitling: subtitlers often keep the original flavour of the source language in order to meet fans' expectations since the majority of fans aim to learn American language and culture by watching videos of American comedy and talk-shows, while the 'domestication' strategy in subtitling is still controversial among Chinese audiences (Hsiao 2014, 84-87). In addition, since the presence of ECs might demand more processing effort, simple words are selected in both the subtitles and ECs to reduce the difficulty of text reading.

All the ECs and subtitles in the video were presented in a single line to exclude the possible impact of different subtitle line numbers and thus to ensure that the ECs were the only variable. Each subtitle was shown for no less than 1 second to guarantee enough reading time, but normally no longer than 5 seconds to prevent viewers from missing any other equally critical information in the video. Each EC coincided with the corresponding dialogue, with its temporal length determined by its word count (1 second per 6 Chinese characters was set as a routine practice[2]). The inserted ECs and their temporal lengths, corresponding contexts and dialogues are presented in the Appendix.

The subtitle quality has been assessed by two professional subtitlers with over 5 years subtitling experience. The assessment was conducted according to the *Target Text Quality Requirements for Translation Services* (2005), and the results met the *Quality Requirements* that an overall error rate of the subtitles was lower than 1.5‰.

4.3. Eye Movement Recording and Analysis

All the participants' eye movements were registered using a Tobii TX300 eye-tracker (300 Hz), a remote tracker which allows for unrestrained head movement. The eye-tracker was

[2] This is a subtitling practice adopted by FabuTrans Company (Beijing, China) in which the subtitler works for, and is stipulated in their internal subtitling manual. The translation company is one of the major translation companies in Beijing.

connected to a 23” LCD monitor serving as the presentation screen. The screen resolution was set at 1280x1024 pixels and the fixation radius was 35 ppi (pixel per inch) (the default setting of the Tobii system). Each participant was asked to sit 60-65 cm away from the eye-tracker. The ambient lighting in the lab was kept relatively constant with the same lighting arrangements for all sessions. The collected eye-tracking data was analysed by Tobii Studio 3.3.0 software supplied by Tobii technology.

The eye-tracking data was designed to provide statistical evidence of the participants’ processing efforts in reading subtitles and ECs. To analyse the results of the test group, three areas of interests (AOIs) were created (see Figure 3): one circled subtitles and was activated when subtitles appeared; a second one circled ECs and was only activated when ECs appeared; and a third AOI circling the whole screen and activating throughout the video was created to calculate the attention devoted to images and texts. 19 segments of video for the 19 ECs and their corresponding subtitle(s) were abstracted from the video. Similar procedures were also applied to the analysis of the results of the control group, except that the AOI for the ECs was deducted.

Figure 3. The setting of the AOIs on the subtitle and the EC

Five types of data were obtained from both AOIs: time to first fixation (TTFF), mean fixation duration, total fixation duration, fixation count and visit count:

(1) The sequence of reading subtitles and ECs can be observed by comparing the TTFF within the AOIs of subtitles and ECs, and checking the dynamic eye-movement recordings. To be precise, if viewers have a shorter TTFF on a subtitle than on a concurrent EC, it means they look at the subtitle ahead of the EC.

(2) Total fixation duration and fixation count can be used to observe how the viewers allocate their attention. The respective amounts of processing effort distributed to subtitles and ECs can be compared through the total fixation duration on each AOI, while the fixation count can reveal whether viewers spend more effort on visiting specific AOIs. The time spent viewing whole screen images can be calculated by using the total fixation duration on the whole screen minus the fixation duration on subtitles and ECs.

(3) A mean fixation duration and visit count was used to investigate the reading patterns of the viewers. If the visit count within the AOI of subtitles is relatively high, it means there are more deflections between the subtitle area and other parts of the screen. A regular reading pattern requires the mean fixation to be as long as possible and the deflections to be as few as possible.

4.4. Questionnaire and retrospective interviews

The questionnaire consisted of two parts. Part A asked about the participants' subjective judgements on subtitles and ECs, as well as their feelings about their comprehension of the video. Both the control and test groups had the first three questions in common. Firstly, all participants were asked whether they had watched the video before; secondly, participants were asked to what extent they though they understood the jokes in the video. The responses to this question were compared with the results in Part B. Thirdly, they were asked how fast they thought the subtitles were, instead of directly asking whether they encountered difficulty in reading the subtitles. Asking about the speed of the subtitles is thought to be more objective and effective in measuring viewers processing efforts (Caffrey 2012).

Three more questions were specifically designed for the test group to obtain feedback on the reception of the ECs. Of these, the last question elicits participants' comments on the ECs in terms of the font size, colour, speed, etc. In addition to solid objective evidence indicating

the usefulness of ECs, EC receivers' subjective opinions and expectations are equally important.

Part B of the questionnaire explored whether participants truly understood the video, or gained some PCEs, by asking six open-ended questions that referred to the video content, especially concerning the jokes and punch lines in the video where loud laughter from the live audience could be heard. It is possible that whether aware or not, some participants still had some misunderstandings, although some declared they almost or even fully understood the video. The number of correct answers for each participant group is a decisive factor in determining whether the inserted ECs improved comprehension of the AV content.

4.5. Procedure

All participants were tested individually in a UK university's eye-tracking laboratory. They were firstly asked to sign a written consent form before formally participating in the research. Then, they were instructed by the researcher to watch the video as they would usually do at home, and afterward were told to answer a questionnaire regarding the video and themselves.

Each experiment started with a 5-point calibration session, followed by a warm-up test involving watching a 30-second video, which aimed to provide participants with some background knowledge about the video in the formal test, and an opportunity to familiarise themselves with the eye-tracking equipment. During the formal test, each participant watched the given video only once. Soon after viewing the video, they were asked to fill in Parts A and B of the questionnaire. Based on their answers, the researcher then carried out retrospective interviews concerning participants' video-watching habits, and elicited comments on subtitles and ECs, and further explanations of some unclear answers.

4.6 Quality assessment of eye-tracking data

Quality assessment on collected eye-tracking data should go beyond data analysis. In this research Hvelplund's (2011; 2014) three criteria were applied to assess the quality of the eye-tracking data: **G**aze **T**ime on the **S**creen as a percentage of total production time (GTS) [(total fixation duration/total task time)*100], **G**aze sample to **F**ixation **P**ercentage (GFP) [Total fixation duration/(total fixation duration + total saccade) *100], and **M**ean **F**ixation **D**uration (MFD). Unlike the task of written translation, the 3.46-minute video-watching task in our experiment requires participants' eye-fixations mostly on the screen, hence, comparatively high thresholds were adopted to guarantee the reliability of the eye-tracking data.

Participant data were considered invalid when the GTS score was lower than one standard deviation below the mean GTS score (73.64%), the GFP score was lower than 85%, or the MFD was shorter than 200ms. As a result of applying two out of the three quality assessment criteria, data for 2 participants (one from the test group and one from the control group) were identified as invalid and were accordingly removed from further analysis. The percentage of invalid data was 5.6%.

5. Results

5.1. Positive Cognitive Effects (PCEs)

The PCEs in the present research are reflected by the participants' self-evaluations of their comprehension of the video, and the recorded accuracy rate for answers to Part B of the questionnaire.

5.1.1. Self-evaluated PCEs

In Part A of the questionnaire, all participants were asked: 'To what extent do you think you understood the jokes in the video?' This question aimed to elicit the participants' PCE based on their self-evaluations.

Table 2. Self-evaluated PCE based on Q2

Group	Total Num.	Not at all	A little bit	Half	Almost	Fully	t-test*
Test	17	1	8	5	2	1	t(17)=0.85 p=0.4>0.05
Control	17	0	8	5	2	2	

* The t-test is calculated not with the numbers of participants choosing each option (as shown in Table 2), but with each participant's choice. That is, a value is assigned to each option, (i.e., 'not at all=1', 'a little bit=2',…, 'almost=5') and then an unpaired t-test is conducted between two groups.

Before the experiment it was expected that the test group, having had the opportunity to access the ECs giving more background knowledge, particularly concerning the understanding of jokes, would be more positive in answering this question. As Table 2 reveals however, the number of responses allocated to each of the 5 categories is nearly the same, with an up-paired t-test result $p>0.05$, suggesting there was no statistically significant difference between the two groups. Interestingly, the only participant from the test group who reported no understanding of the jokes at all, correctly answered 4 out of the 6 questions. In contrast, the only participant from the control group who claimed to have understood all the

jokes turned out to have misunderstood them all. This mismatch between the self-evaluated PCEs and the tested PCEs will be further discussed later.

When asked if the provision of ECs is helpful to them in understanding the video, 16 out of 17 (94.1%) participants in the test group gave a positive reply. The only participant who declared that the ECs were not helpful at all, however, correctly answered 4 out of the 6 questions. When asked for the reasons in his post-experiment interview, he mentioned that the jokes had been clearly explained by the ECs. Thus, it can be assumed that, whether he admitted it or not, the ECs did help him to understand the jokes. Conversely, the retrospective interview data shows that the majority of participants (88.2%) in the control group, based on their personal experience, also affirmed the necessity for, and positive effects of, ECs in the reception of foreign language AV products.

5.1.2 Tested PCEs

Questionnaire Part B included 6 questions relevant to the jokes being explained by the ECs. Both groups were asked to answer these questions. As shown in Table 3, the test group had an average of 3.24 (SD=1.6) correct answers, much higher than that the control group with 0.24 (SD=0.56). The un-paired t-test result was $p<0.05$, suggesting a statistically significant differences between the two groups.

Table 3. The number of correct answers by each participant

	PS								**TS**						**UL**			**Avg. (SD)**	**t-test**
Participant No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		
Test	4	4	1	5	3	3	3	1	6	3	5	4	4	5	2	1	1	3.24 (1.60)	t(17)=7.29, p=0.00<0.05
Control	0	0	0	0	0	0	0	0	0	1	0	1	2	1	0	0	0	0.24 (0.56)	

Apart from comparing the test results between the two groups, we also calculated the percentage of correct scores for each question. In Table 4, both Q2 and Q3 have been correctly answered by 14 participants, gaining the highest correct response rate (41.4%) of all the 6 questions. By contrast, Q6 was correctly answered by only 3 participants from the control group, gaining the lowest correct response rate (8.82%). Participants under test conditions achieved the highest correct response rate for Q2 (82.35%), while participants from the control group achieved the highest correct response rate for Q1 (17.64%). No participant in the control group correctly understood the allusions that Q2, Q5 and Q6 refer to.

The relationship between particular questions and their relevant ECs is further considered in Section 6.

Table 4. Number of correct answers for each question

Question	Test				Control				Overall
	PS	TS	UL	Subtotal	PS	TS	UL	Subtotal	
Q1	3	5	0	**8**	0	3	0	**3**	**11**
Q2	7	6	1	**14**	0	0	0	**0**	**14**
Q3	6	6	0	**12**	0	2	0	**2**	**14**
Q4	4	3	1	**8**	0	2	0	**2**	**10**
Q5	6	3	2	**11**	0	0	0	**0**	**11**
Q6	0	3	0	**3**	0	0	0	**0**	**3**

5.2. Processing Efforts

In this research, the processing efforts were examined and measured through the triangulation of questionnaire, retrospective interviews and eye-tracking data. The combination of subjective and objective data was aimed to accurately reveal the processing efforts made by the two groups.

5.2.1. Questionnaire and retrospection results

Processing efforts were partly measured by inviting participants' self-evaluations of the subtitling speed. All the control group participants felt that the speed of the subtitles in the video was 'normal'. Among the test group, while the majority of the participants (70.59%) also thought the speed of the subtitles was 'normal', 4 participants rated it 'a little too fast' and 1 found the subtitles to be 'a little bit slow'. In addition, among the 12 test participants who thought the subtitle speed was 'normal', 2 complained about the short duration of the ECs. In fact, there were altogether 6 test group participants (35.29%) claiming that they have difficulties in finishing reading both the subtitles and the ECs. At the same time, about one third of the test group participants suggested that the ECs should be put at the bottom to reduce the cognitive effort involved in moving the eyes between the top and the bottom of the screen. A further 3 test group participants mentioned that the similarity in colour between some the ECs and the background colour led to increased reading difficulty.

Table 5. Feedback on the speed of the subtitles (Q3)

Group	Total Num.	Too slow	A little bit slow	Normal	A little bit fast	Too fast
Test	17	0	2	11	4	0
Control	17	0	0	17	0	0

Table 6. Feedback on the impact of provision of ECs on comprehension of the video (Q4)

Group	Total Num.	No help, Increased PE	No help, No increased PE	Helpful, Increased PE	Helpful, No increased PE
Test	17	1	1	10	5
Percent	/	5.88%	5.88%	58.82%	29.41%

In the questionnaire, the test group was also asked about their feelings concerning the impact of the provision of ECs on comprehension of the video content. As can be seen from Table 6, 88.24% participants confirmed that the provision of EC had a positive effect on their comprehension of the video. At the same time, 64.71% participants believed that the existence of the ECs did increase their processing effort; for example, they had been distracted by the ECs at some points and thus missed other details shown in the video.

5.2.2. Eye-tracking results

Based on the research mentioned earlier (d'Ydewalle & Gielen 1992; de Linde & Kay 1999; Kruger et al. 2013), it was hypothesised that, first, the test group would spend more time in reading the texts than the control group; and second, the presence of the ECs would reduce the reading time spent on the subtitles and increase the deflections, which indicates a less regular reading pattern.

Table 7. Eye-tracking data on reading the subtitles, ECs and image

	Total fixation duration (s)					**Fixation count**			
	Control (Subtitle)	**Control (Image)**	**Test (Subtitle)**	**Test (Subtitle+EC)**	**Test (Image)**	**Control (Subtitle)**	**Control (Image)**	**Test (Subtitle)**	**Test (Image)**
Avg.	58.30	143.57	48.66	82.24	107.92	257.65	465.18	189.71	334.41
t-test	Control(Subtitle) vs. Test (Subtitle) t(17)=1.09, p=0.28>0.05 Control(Subtitle) vs. Test (Subtitle+EC) t(17)=-2.62, p=0.01<0.05 Control (Image) vs. Test (Image) t(17)=4.33, p=0.000<0.05					Control(Subtitle) vs. Test (Subtitle) t(17)=2.25, p=0.03<0.05 Control (Image) vs. Test (Image) t(17)=4.93, p=0.000<0.05			
	Mean fixation duration (ms)					**Average visit count**			
	Control (Subtitle)	**Control (Image)**	**Test (Subtitle)**	**Test (EC)**	**Test (Image)**	**Control (Subtitle)**		**Test (Subtitle)**	
Avg.	223.53	314.49	246.47	270.59	328.51	90.53		78.65	

t-test	Control(Subtitle) vs. Test (Subtitle) t(17)=-1.61, p=0.12>0.05	Test (Subtitle) vs. Test (EC) t(17)=-1.58, p=0.12>0.05	t(17)=2.01, p=0.027<0.05
	Control (Subtitle) vs. Control (Image) t(17)=-4.79, p=0.000<0.05	Test (Subtitle) vs Test (Image) t(17)=-4.51, p=0.000<0.05	
	Control(Subtitle) vs. Test (EC) t(17)=-3.05, p=0.005<0.05	Control (Image) vs. Test (Image) t(17)=-0.63, p=0.53>0.05	

As can be seen from Table 7, the test group had a higher average total fixation duration on the texts (subtitle + EC = 82.24s) compared to the control group (subtitle = 58.30s). The un-paired t-test p<0.05, suggests that the difference is statistically significant. This result therefore supports our first hypothesis that more time was devoted to read texts by the test group than by the control group, and also suggests that there is a remarkable increase in the processing effort with the presence of the ECs, a finding that is also supported by the questionnaire results on processing efforts in 5.2.1.

The eye-tracking results presented in Table 7 however, do not support our second hypothesis that less time will be spent for reading subtitles and the reading of subtitles will be less regular. Compared with the control group, the test group did not experience an obvious decrease in their total fixation duration in the AOI for subtitles (48.66s vs 58.30s, p>0.05). Although the fixation count on subtitles is lower in the test group (189.71 vs 257.65, p<0.05) with statistical significance, the mean fixation duration on subtitles between test and control groups showed no statistically significant difference (246.47ms vs 223.53ms, p>0.05). The fact that the test group spent significantly less time than the control group in viewing images (107.92s vs 143.57s, p<0.05), indicates that they sacrificed their time on image-observation when ECs were presented. This finding is in line with Hegarty's (1992) research findings that suggest text tends to be prioritised over other semiotic modes in a multimodal context. Moreover, the average visit counts on subtitles did show a statistically significant decrease (78.65 vs 90.53, p<0.05) in the presence of ECs, which could indicate that there were less deflections for the test group in their reading pattern.

Another unexpected result is that, while the mean fixation durations on subtitles and on ECs by the test groups demonstrate no statistically significant difference, the mean fixation duration on ECs by the test group is significantly higher than that on subtitle by the control group (270.59ms vs 223.53ms, p<0.05). This echoes the research of Rayner & Pollatsek (1989) that fixation duration becomes longer as texts becomes conceptually more difficult. In the current research, subtitles are simply texts spoken in the video, which are direct and

explicit and thus, in relevance-theoretic terms, more relevant; while ECs are indirect and implicit, and thus require more cognitive effort to build up the relevance. More cognitive effort required indicates higher difficulty of processing and thus the mean fixation duration on ECs is longer.

The mean fixation duration on images by both groups is between 310ms and 330ms, which is longer than that on texts with statistically significant difference. This is in line with previous research reports that looking at scenes require longer fixations than reading texts (Szarkowska & Kruger 2015) and that the mean fixation duration of perceiving scenes is about 330ms, longer than that of reading subtitles (Rayner 1998).

The eye-tracking data also provide evidence for the reading order of ECs and subtitles. Among the 10 video clips where ECs appear on the screen simultaneously with their corresponding subtitles, 8 have a shorter average TTFF on subtitles than on ECs. It means when a subtitle and an EC appear on the screen simultaneously, the viewers tend to look at the subtitle first, and then at the EC.

6. Discussion

The current research is aimed at investigating whether the provision of ECs improves the comprehension of AV content by examining the relevance of ECs to audiences. As illustrated above, the test group spent a greater amount of processing effort to read the texts in the video, however, the presence of ECs does not necessarily influence the processing efforts devoted to reading subtitles, as the test and control groups are similar in their total fixation durations on subtitles. These results are in line with d'Ydewalle & Gielen's (1992) finding that subtitle processing is almost automatic and effortless. The fact that test group has statistically shorter total fixation duration on images than the control group indicates that, instead of distributing an unchanged amount of processing effort to both subtitles and ECs, the participants actually sacrificed processing efforts devoted to images and invested additional effort on processing the ECs. It is worth noting however, that the lower fixation count, the almost unchanged mean fixation duration and fewer deflections among the test group suggests their reading pattern on subtitles was more regular, which could balance out the increased effort on text reading (de Linde & Kay 1999; Lautenbacher 2013, 148).

At the same time, the test group had a higher average correct response rate than the control group, revealing an increase in PCEs among the test group. The fact that most PS and TS participants reported that they would stop the video to read the ECs which are appealed to them when watching AV products at home, and that some participants in the control group

affirmed the value of ECs in some specific parts of AV products, clearly supports the argument that the increased processing effort invested in ECs is acceptable to young audiences.

Although the general tendency in changes in processing effort and PCEs accords with our prior expectation that the ECs would improve the comprehension of the video, the fact that only 3 participants correctly answered Q6 might pose a challenge to this conclusion. Q6 asks why John Boehner mentions getting a 'Grand Bargain' on a Chevy Tahoe when talking to President Obama about life after retirement. The 'Grand Bargain' in this context refers to the opportunity to obtain a Chevy Tahoe at a lower price. It also alludes to an attempted compromise in the 2011 budget debate between the Democrats and the Republicans, represented by President Obama and Boehner respectively, which ended up in failure. The 'Grand Bargain', as a famous political event in American society, is often translated as '*大谈判*' (gloss: "Grand Negotiation") or '*大交易*' (gloss: "Grand Deal"). As illustrated in Table 8, EC18 was inserted to explain this allusion. 'Grand Bargain' in the subtitle was translated as '*高折扣*' (gloss: "high discount rate"), but left untranslated in EC18, assuming the viewers would notice the pun.

Table 8. The subtitle and EC regarding Q6

Context	Spoken dialogue	Subtitle	EC
John Boehner was giving suggestions to President Obama for the life after retirement.	John Boehner: I finally got the grand bargain on a sweet Chevy Tahoe.	*我终于拿到雪佛兰太浩的高折扣了* ("I finally got a high discount rate on a Chevy Tahoe.")	EC18: Grand Bargain *指二人曾就联邦预算问题进行的大谈判 最终以失败收场* ("'Grand Bargain' refers to the negotiation between these two on the federal government budget, which ended in failure.")

The failure of EC18 can be partly attributed to its lack of relevance. Low frequency words, such as '*联邦预算*' (gloss: "federal government budget") and '*大谈判*' (gloss: "Grand Bargain"), normally require more processing effort (Moran 2012). They are unfamiliar to native Chinese speakers not only in their literal sense but also from the point of view of background knowledge. Three participants in the test group mentioned that EC18 "seems to require some background knowledge to understand it". When these words are used in their metaphorical sense, they impose even higher cultural difference-based cognitive loads (Zheng & Xiang 2014), or processing effort. In relevance-theoretic terms, other things

being equal, low-frequency words and concepts require more processing effort than high-frequency ones, as they are less relevant to audiences (Wilson & Sperber 2002, 252).

Similarly, Q4 which is related to EC12, asks about the identity of John Boehner, who is the former Speaker of the House of Representative ("*前众议院议长*"). As most of our participants in the test group were unfamiliar with American official titles, they even forgot the name of the title although it appeared in EC12. Instead, they gave answers with some similar titles based on their vague recollection, such as '*前参议院议长*' (gloss: "former leader of the United States Senate") and '*前议员*' (gloss: "former senator/congressman"). By contrast, Q5, which is also related to EC12, asks why the audience laughed at John Boehner when he wiped his eyes. Half of the participants correctly answered '*爱哭*' (gloss: "cry-baby"), a high-frequency word in Chinese used in EC12.

Table 9. Eye-tracking results from reading EC18 and its corresponding subtitles

	EC18: Total fixation duration (s)		**EC6&7: Total fixation duration (s)**		
	Control (Subtitle)	**Test (Subtitle+EC)**	**Control (Subtitle)**	**Test (Subtitle)**	**Test (Subtitle+EC)**
Avg.	2.43	2.91	1.3	1.08	2.47
t-test	t(17)=-1.63, p=0.11>0.05		Control (Subtitle) vs. Test(Subtitle) t(17)=0.88, p=0.38>0.05 Control (Subtitle) vs. Test (Subtitle+EC) t(17)=-4.16, p=0.00<0.05		

Another reason for the failure of EC18 can be attributed to the limited time frame. As mentioned above, EC18 and its corresponding subtitle appear simultaneously, leaving participants to read 39 characters in 4.5 seconds. With such an intensive reading task, the total fixation duration of the test group on the subtitle and EC combined (2.91s) is not much longer than that of the control group on the subtitle alone (2.43s), with the t-test result of p>0.05, indicating no statistically significant difference. In other words, the test participants did not manage to spend significantly more processing time than the control participants. Our interview data show that 6 test participants did not manage to finish reading both the subtitle and the EC.

Unlike EC18, EC6 & 7 gave the test participants sufficient time to read both the subtitles and the ECs. As EC6 & 7 appeared about 1 second later than the corresponding subtitles, together, EC6 & 7 and the subtitles allowed participants to read 34 Chinese characters in nearly 6 seconds. The reading task for this segment was less demanding than that for EC18.

Consequently, the total fixation duration of the test group (Subtitle + EC) was much longer than that of the control group (Subtitle only) (2.47s vs 1.3s), with a t-test significance of $p<0.05$, indicating statistically significant difference. This indicates that the test participants had obviously devoted additional time and effort to reading the additional text, i.e., the EC. The additional effort resulted in the highest correct rate in Q2, directly related to EC6 & 7.

Table 10. The time to first fixation (TTFF) on EC18 and the corresponding subtitle in the test group (s)

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Avg.	t-test
Subtitle	0.34	0.52	0.32	0	0	0.12	0.06	0.24	1.03	0.41	0.53	2.73	**0.53**	0.46	0.97	0.48	**0.9**	0.57	t(17)=-3.92, p=0.00<0.05
EC	1.86	2.64	1.88	0.35	0.22	0.61	1.49	0.88	2.08	2.78	2.15	2.15	**0.05**	0.94	1.95	2.39	**0.5**	1.58	

As can be seen from Table 10, most test group participants had a shorter TTFF on the subtitle than for EC18, meaning that they generally first fixated on the subtitle. In other words, they tended to read the subtitle before EC18. This reading pattern exposed by the eye-tracking data is echoed in participants' retrospective data, where 88.23% of participants answered that they read the subtitles before the ECs if both appeared simultaneously on the screen. Only 2 participants had a longer TTFF on the subtitle, suggesting they read EC18 before the corresponding subtitle. The retrospective interview data suggest that they give particular attention to the EC if it appears together with a subtitle. Participant No.13 even explicitly explained that she could easily understand the dialogues by listening to them without looking at the subtitles, and on many occasions, the ECs were very helpful to her in better comprehending the AV content.

The reading pattern in which subtitles are usually given a priority position compared to the ECs reflects the basic mechanism of human cognition. Human beings always subconsciously distinguish the relevant (or more relevant) information from the irrelevant (or less relevant) information, and first of all retrieve explicit information contained in the stimulus, and then retrieve implicit information (Wilson & Sperber 2002; Braun 2016). Since the dialogues presented by the subtitles are usually recognised as more explicit and direct messages, they are deemed to be more relevant information and retrieved in the first instance by most of the audience. ECs, however, mostly explain implicit or implied meanings, such as wordplays, allusions, idioms and figurative language, and hence are given secondary position and processed later.

7. Conclusion

Overall, the test group's significant increase in their average number of correct answers to Part B of the questionnaire indicates that these participants experienced more PCEs. It is safe to conclude therefore, that the ECs made a positive contribution to comprehending culturally specific allusions in the video. Moreover, most participants believe that ECs help them to understand the video, and they reported that they would pause to read the ECs when necessary, as they wanted to learn these culturally-specific allusions.

In terms of processing effort, both the questionnaire and eye-tracking results suggest that the test participants devoted more processing effort to reading texts by reducing the time spent on viewing images. Nearly 2/3 of the test participants reported that the presence of ECs incur distraction from other parts of the video. Unlike the previous assumption that the ECs would increase the irregularity of text reading in the subtitled video however, the two groups had a similar total fixation duration in the AOI of subtitles, which means the presence of ECs does not necessarily reduce the processing time allocated to subtitles.

Another unexpected result was that the mean fixation duration of the test group remained almost unchanged, while the visit counts in the AOI of subtitles declined, representing a more regular reading pattern. Such a pattern is beneficial in information processing, and helps the test group to better comprehend the AV content.

These findings also offer some suggestions on the making of AV subtitles and ECs, in terms of timing and translation strategy. Since the subtitles are often read before the ECs and other elements of the video, an appropriate time delay before showing the EC might be helpful in reducing the pressure over reading both the subtitle and the EC in a limited time frame, also, high-frequency words and concepts, which are of greater relevance to foreign language audiences, tend to be received more easily.

There are some limitations to the present research although there are some interesting findings to report. Firstly, the participants are not sufficiently representative of the online video users in China. Most participants in this research were in their 20s and 40s with good educational background, and more samples representing a wider group of people should be considered in future research. Secondly, the number of the questions interrogating the participants' understanding of the video is rather limited: only 6 questions were devoted to determining the level of PCE experienced by the viewers. The results of such a restricted experimental sample could possibly be affected by other factors. Thirdly, in this research only a political TV show was selected as the stimulus: a wider selection of AV products

covering different topics, and in different styles, would increase the generalisability of the research.

References

Bisson, Marie-Josée., Walter J. B. Van Heuven, Kathy Conklin and Richard J. Tunney (2014), "Processing of Native and Foreign Language Subtitles in Films: An Eye Tracking Study", *Applied Psycholinguistics* 35 (2): 399-418.

Braun, Sabine. 2016. "The Importance of Being Relevant? A Cognitive-pragmatic Framework for Conceptualising Audiovisual Translation." *Target* 28(2): 302-313.

Caffery, Colm. 2012. "Using an Eye–tracking Tool to Measure the Effects of Experimental Subtitling Procedures on Viewer Perception of Subtitled AV Content." In *Eye-tracking in Audiovisual Translation*, edited by Elisa Perego, 135-156. Roma: Aracne.

Dwyer, Tessa. 2015. "From Subtitles to SMS: Eye Tracking, Texting and *Sherlock*", *Refractory: A Journal of Entertainment Media*, 25. Accessed 11 Nov 2017: http://refractory.unimelb.edu.au/2015/02/07/dwyer/

d'Ydewalle, Géry, and Ingrid Gielen. 1992. "Attention Allocation with Overlapping Sound, Image, and Text." In *Eye Movements and Visual Cognition: Scene Perception and Reading*, edited by Keith Rayner, 415–427. New York: Springer.

d'Ydewalle, Géry, and Wim De Bruycker. 2007. "Eye Movements of Children and Adults while Reading Television Subtitles." *European Psychologist* 12 (3): 196-205.

d'Ydewalle, Géry, Johan van Rensbergen, and Joris Pollet. 1987. "Reading a Message When the Same Message is Available Auditorily in Another Language: The Case of Subtitling." In *Eye Movements: From Physiology to Cognition*, edited by J. K. O'Regan and A. Lévy–Schoen, 313–321. Amsterdam: North– Holland.

de Linde, Zoe, and Neil M. Kay. 1999. *The Semiotics of Subtitling*. Manchester: St. Jerome.

Desilla, Louisa. 2012. "Implicatures in Film: Construal and Functions in Bridget Jones Romantic Comedies." *Journal of Pragmatics* 44: 30-53.

Ghia, Elisa. 2012. "The Impact of Translation Strategies on Subtitle Reading." In *Eye-tracking in Audiovisual Translation*, edited by Elisa Perego, 157-182. Roma: Aracne.

Hegarty, Mary. 1992. "The Mechanics of Comprehension and Comprehension of Mechanics." In *Eye Movements and Visual Cognition: Scene Perception and Reading*, edited by Keith Rayner, 428–443. New York: Springer.

Hsiao, Chi-hua. 2014. *The Cultural Translation of U.S. Television Programs and Movies: Subtitle Groups as Cultural Brokers in China*. PhD Thesis. Los Angeles: University of California.

Hvelplund, Kristian Tangsgaard. 2011. *Allocation of Cognitive Resources in Translation: An Eye-tracking and Key-logging Study*. PhD thesis. Copenhagen: Copenhagen Business School.

Hvelplund, Kristian Tangsgaard. 2014. "Eye Tracking and the Translation Process: Reflections on the Analysis and Interpretation of Eye-tracking Data. *MonTI Special Issue – Minding Translation* 201-223.

Kruger, Jan-Louis, Esté Hefer, and Gordon Matthew. 2013. "Measuring the Impact of Subtitles on Cognitive Load: Eye Tracking and Dynamic Audiovisual Texts." *Proceedings of the 2013 Conference on Eye Tracking South Africa*, 62–66. New York: ACM. DOI: 10.1145/2509315.2509331

Lautenbacher, Olli Philippe. 2012. "From Still Pictures to Moving Pictures." In *Eye-tracking in Audiovisual Translation*, edited by Elisa Perego, 135-155. Roma: Aracne.

Moran, Siobhan. 2012. "The Effect of Linguistic Variation on Subtitle Reception." In *Eye-tracking in Audiovisual Translation*, edited by Elisa Perego, 183-222. Roma: Aracne.

National Standard of the People's Republic of China. 2005. *Target Text Quality Requirements for Translation Services*. Beijing: Standards Press of China.

Perego, Elisa. 2012. *Eye Tracking in Audiovisual Translation*. Roma: Aracne.

Rayner, Keith. 1998. "Eye Movements in Reading and Information Processing: 20 Years of Research." *Psychological Bulletin* 124(3): 372–422.

Rayner, Keith., and Alexander Pollatsek. (1989). *The Psychology of Reading*. Englewood Cliffs, NJ: Prentice Hall.

Sperber, Dan, and Deirdre Wilson. 1995. *Relevance: Communication and Cognition*. 2nd ed. Oxford: Blackwell.

Szarkowska, Agnieszka and Jan Louis Kruger. 2015. "Subtitles on the Moving Image: An Overview of Eye Tracking Studies." *Refractory: a Journal of Entertainment Media*, 25. Accessed 11 Nov 2017: http://refractory.unimelb.edu.au/2015/02/07/kruger-szarkowska-krejtz/

Tannen, Deborah. 1989. *Talking Voices: Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.

Thomas, Jenny. 1995. *Meaning in Interaction: An Introduction to Pragmatics*. London: Longman.

Wilson, Deirdre, and Dan Sperber. 2002. "Relevance Theory." *UCL Working papers in Linguistics* 14: 249–287.

Yus, Francisco. 2008. “Alterations of Relevance in Cyber–Media.” *Universitas Psychologica* 7(3): 629–642.

Zheng, Binghan, and Xia Xiang. 2014. “The Impact of Cultural Background Knowledge in the Processing of Metaphorical Expressions.” *Translation and Interpreting Studies* 9(1): 5-24.

Appendix: ECs inserted in the video

EC No.	Time	Context	Dialogue	EC
1	00:05.20-00:03.76	Joe Biden was cleaning his sunglasses while talking to President Obama.	**Joe Biden:** Which one do you like better? This? Or this? **Obama:** Joe, they are the same.	乔·拜登出了名地热爱雷朋飞行员墨镜 Joe Biden is a famous fan of Ray-Ban Aviator Sunglasses.
2	00:03.76-00:05.80			经常戴该款墨镜出席活动 (He) Often wears this type of sunglasses to attend important events.
3	00:33.05-00:35.03	President Obama went to DMV.		欢迎光临华盛顿特区车管局 Welcome to the Washington D.C. Department of Motor Vehicles
4	00:39.14-00:40.55	President Obama was waiting at the DMV and his number is 44.	**DMV staff:** 44!	奥巴马是美国第 44 任总统 President Obama is the 44th president of the USA.
5	00:41.21-00:44.12	After waiting for a long time, President Obama finally got called.	**President Obama:** Finally…	吐槽美国车管局常被批评办事效率低下 (This is) reflecting the fact that the DMV in the USA is often criticised for its low efficiency.
6	00:47.23-00:48.93	When hearing Hussein as the middle name of President Obama, the DMV staff presented an alarmed and scared expression.	**DMV staff:** What’s the name? **Obama:** Barack Hussein Obama. **DMV staff:** Yikes.	侯赛因是阿拉伯名 Hussein is an Arabic name.
7	00:48.93-00:51.23			奥巴马曾因此被怀疑与穆斯林有关联 President Obama was thus suspected over his relations with Muslims.
8	00:54.47-00:58.27	The DMV staff asked Obama to illustrate his birth certificate.	**DMV staff:** Well, since you don’t have a driver’s license, you’re gonna need a birth certificate.	特朗普曾质疑奥巴马生于肯尼亚 没资格做总统 Donald Trump once suspected that Obama was born in Kenya, meaning he was not qualified to be US president.
9	01:00.54-01:04.84	The DMV staff kept questioning the validity of President Obama’s birth certificate.	**Obama:** It’s real. **DMV staff:** Is it? **Obama:** It’s real. **DMV staff:** But is it?	奥巴马为此曾公开出生证明 但仍被怀疑系伪造 Obama released his birth certificate online, but some still doubt its validity.
10	01:29.31-01:20.70	News title.		突发新闻 Breaking news

11	01:20.91-01:21.99	Wolf Blitzer broadcast the news about Michelle Obama posted a Snapchat.		CNN 新闻主播沃尔夫·布利策 (This is) the CNN news broadcast Wolf Blitzer.
12	02:03.55-02:06.36	President Obama and John Boehner were sitting in the cinema at the end of *The Toy Story III*, and John Boehner was wiping away his tears.		前众议院议长约翰·博纳 常被吐槽爱哭 (This is) John Boehner, the former House of Representative Speaker, who is often mockingly criticised as a 'cry-baby'
13	02:14.51-02:16.51	John Boehner was obviously surprised when President Obama asked John Boehner for suggestions on dealing with retirement.	**Obama:** So, you got any advice for me? **John Boehner:** Now you want my advice?	博纳身为共和党人 与奥巴马争斗多年 As a republican, Boehner has been in conflict with Obama for years.
14	02:32.10-03:34.06	John Boehner suggested Obama to be himself.	**Obama:** I can wear my mom jeans in peace. I hate these tight jeans.	媒体曾吐槽奥巴马穿老妈牛仔裤 The media often made fun of Obama's jeans as mom jeans.
15	02:46.38-02:49.27	John Boehner gave Obama some suggestions.	**John Boehner:** … And you know McDonald's now serves breakfast all day long. **Obama:** And Michelle's gonna to be at spin class, so she'll never know.	米歇尔曾倡议禁止校园内贩卖垃圾食品 Michelle Obama proposed a prohibition on selling junk food on campus.
16	02:54.04-03:56.99	John Boehner persuaded President Obama to relax and then leave the office hum the theme tune.	**John Boehner:** Right, let it go. And it won't be long, you'll able to walk right out of your office, singing 'zip-a-dee-doo-dah, zip-a-dee-ay.'	博纳退休前上班时常哼的迪士尼电影《南方之歌》的主题调 Before retiring, Boehner often hummed the theme tune in the Disney film *Song of the South*.
17	02:57.00-02:58.34			因出现在他的退休纪录片中而众所周知 (This) became well known because it appeared in Boehner's retirement documentary.
18	03:03.46-03:08.17		**John Boehner:** I finally got the grand bargain on a sweet Chevy Tahoe.	Grand bargain 指二人曾就联邦预算问题进行大谈判 最终失败 'Grand bargain' refers to the negotiation between this two on the federal government budget, which ended up in failure.
19	03:29.26-03:31.74		**Wolf Blitzer**: Breaking news, former President Barack Obama on his 347th round of Golf for the year, and it's totally great. And Gloria, not a problem for anybody?	奥巴马爱好高尔夫 为此曾受批评 Obama is a big fan of golf, for which he was once criticised.

Authors' biosketch:

Dr Binghan ZHENG is an Associate Professor in Translation Studies at Durham University, and a by-fellow of Churchill College, Cambridge University. His research focuses on empirical studies of translation and interpreting. He is the author of monograph *Choice-making in the process of English-to-Chinese translation: An empirical study*, and has published widely in journals such as *Across Languages and Cultures*, *Babel*, *Perspectives, Translation and Interpreting Studies, Foreign Language Teaching & Research, Journal of Foreign Languages*. He has edited a special issue of *Translation and Interpreting Studies* entitled "Towards a Comparative Translation and Interpreting Studies" (2017).

Ms Mingqing XIE obtained her MA (Distinction) from the School of Modern Languages and Cultures, Durham University. She has worked as a freelance translator and subtitler for over 5 years, and is now working at Beijing Jiaotong University.